

МІНІСТЕРСТВО ОСВІТИ І НАУКИ УКРАЇНИ  
НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ  
"КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ"  
ФІЗИКО-ТЕХНІЧНИЙ ІНСТИТУТ

КРИПТОГРАФІЯ  
Комп'ютерний практикум

Робота №1

Виконали:  
Сернова А.Р., Колесник А.М.  
студенти групи ФБ-93

Перевірила:  
Селюх П.В.

Київ-2021

**Мета роботи:** Засвоєння понять ентропії на символ джерела та його надлишковості, вивчення та порівняння різних моделей джерела відкритого тексту для наближеного визначення ентропії, набуття практичних навичок щодо оцінки ентропії на символ джерела

**Постановка задачі:** реалізувати програму, що підраховує значення ентропії для літер та біграм (з перетином та без) у текстовому файлі, як з урахуванням пробілів, так і без них.

Під час виконання лаб. роботи виникли деякі труднощі: найбільша це вибір іншої мови програмування (вона була обрана через те, що python – дуже багата на функції з рядками мови), також деякі труднощі з фільтруванням тексту, які ми вирішили за допомогою різних інтернет ресурсів таких як [stackoverflow](#) та подібних. Дещо виникали складності під час початку роботи з біграммами, а саме продумування варіантів та алгоритмів їх підрахунку.

Нижче наведені таблиці для літер російського алфавіту без «ё» і «ъ» (вони були замінені) без урахування пробілів (*таб. 1*), та з пробілами (*таб. 2*)

о	0.1101
е	0.0841
а	0.0790
т	0.0661
и	0.0657
н	0.0629
с	0.0552
л	0.0454
р	0.0430
в	0.0410
к	0.0388
м	0.0319
у	0.0292
п	0.0288
д	0.0287
я	0.0240
ы	0.0196
ь	0.0193
г	0.0176
ч	0.0173
з	0.0168
б	0.0163
й	0.0104
ж	0.0100
х	0.0094
ш	0.0091
ю	0.0069
щ	0.0041
э	0.0034

_	0.1644
о	0.0920
е	0.0703
а	0.0660
т	0.0552
и	0.0549
н	0.0526
с	0.0462
л	0.0378
р	0.0359
в	0.0343
к	0.0324
м	0.0267
у	0.0244
п	0.02404
д	0.02402
я	0.0200
ы	0.0163
ь	0.0162
г	0.0147
ч	0.0145
з	0.0141
б	0.0136
й	0.0087
ж	0.0084
х	0.0079
ш	0.0076
ю	0.0058
щ	0.0035

ц	0.0028
ф	0.0022

Таб. 1

э	0.0028
ц	0.0023
ф	0.0018

Таб. 2

H = 4.475

H = 4.384

Для біграм без урахування пробілів наведено фрагменти з отриманого файлу after.txt: з перетинами (таб. 3) та без (таб. 4)

таб. 3

```

Bigram cross, no spaces
аа -> 0.0002151421081385549  аб -> 0.0016787860564197474  ав -> 0.004372297489020553  аг -> 0.0010909568318207036  ад
ба -> 0.00156020379209141  бб -> 2.1175404344345956e-05  бв -> 4.150379251491807e-05  бг -> 2.2022420518119794e-05  б
ва -> 0.006969249077811141  вб -> 0.00015076887893174321  вв -> 0.00034219453420463066  вг -> 0.00024732872274196074
га -> 0.0014467036248057158  гб -> 5.336201894775181e-05  гв -> 9.063073059380069e-05  гг -> 8.470161737738383e-06  г
да -> 0.004650965810192146  дб -> 0.00011265315111192048  дв -> 0.0012324085328409345  дг -> 6.522024538058554e-05  д
еа -> 0.00021683614048610258  еб -> 0.002485145453852441  ев -> 0.003449049859607069  ег -> 0.004447681928486424  ед
жа -> 0.0016034016169538757  жб -> 1.439927495415525e-05  жв -> 1.6093307301702927e-05  жг -> 2.795153373453666e-05
за -> 0.006252673394798474  зб -> 0.00020074283318439966  зв -> 0.000874120691334601  зг -> 0.0004379073618410744  зд
иа -> 0.00037861622967690566  иб -> 0.0011341546566831693  ив -> 0.00577072119192116  иг -> 0.0010867217509518345  ид
йа -> 0.0001439927495415525  йб -> 0.0002447876742206392  йв -> 0.0007131876183175717  йг -> 0.00022530630222384097
ка -> 0.009034274509471758  кб -> 0.0003498176797685952  кв -> 0.0006013814833794251  кг -> 0.00011011210259059897  к
ла -> 0.004798346624428794  лб -> 0.0002591869491747945  лв -> 0.0006632136640649153  лг -> 0.00038200429437200103  л
ма -> 0.003328773562931184  мб -> 0.0003692990517653935  мв -> 0.0008080534297802417  мг -> 0.00024902275508950845  м
на -> 0.011203482930506558  нб -> 0.00018549654205647057  нв -> 0.0004167319574967284  нг -> 0.00020158984935817349
оа -> 0.00021260105961723338  об -> 0.005327731733037442  ов -> 0.008653964247447306  ог -> 0.0056047060218614875  од
па -> 0.0020963650300902495  пб -> 3.3880646950953528e-06  пв -> 6.7761293901907056e-06  пг -> 4.235080868869191e-06
ра -> 0.008686997878224485  рб -> 0.000534467205651292  рв -> 0.0006954002786683212  рг -> 0.00018803759057779208  рд
са -> 0.002193771890074241  сб -> 0.00020582493022704268  св -> 0.0018134616280497877  сг -> 0.00017872041266627986
та -> 0.007836593639755552  тб -> 0.0002947616284732957  тв -> 0.0028510564409227395  тг -> 0.0003167840489914155  тд
уа -> 0.00017448533179741067  уб -> 0.000837698995862326  ув -> 0.001283229503267365  уг -> 0.00148736040114686  уд
фа -> 0.00029137356377820033  фб -> 4.235080868869191e-06  фв -> 3.388064695095353e-05  фг -> 3.3880646950953528e-06
ха -> 0.0006327210818090572  хб -> 0.00016008605684325543  хв -> 0.0006632136640649153  хг -> 9.740685998399139e-05
ца -> 0.0006428852758943432  цб -> 3.3880646950953528e-06  цв -> 0.00016770920240721996  цг -> 1.0164194085286059e-05
ча -> 0.0023064250411861617  чб -> 7.623145563964544e-06  чв -> 3.133959842963201e-05  чг -> 1.1858226432833735e-05
ша -> 0.0011680353036341228  шб -> 8.470161737738382e-07  шв -> 6.860831007568089e-05  шг -> 8.470161737738382e-07  ш
ща -> 0.0005488664806054471  щб -> 8.470161737738382e-07  щв -> 0.0  щг -> 0.0  щд -> 0.0  ще -> 0.001770263803187321
ья -> 6.352621303303787e-05  ьб -> 0.00042435510306069297  ьв -> 0.002109917288870631  ьг -> 0.00034219453420463066
ьа -> 0.00016008605684325543  ьб -> 0.00032101912986028466  ьв -> 0.0011138262685125972  ьг -> 0.0002219182375287456
эа -> 0.0  эб -> 0.0  эв -> 1.6940323475476764e-06  эг -> 1.3552258780381411e-05  эд -> 6.7761293901907056e-06  эе ->
юа -> 3.9809760167370394e-05  юб -> 0.0005166798660020413  юв -> 0.00022784735074516248  юг -> 0.00011773524815456352
яа -> 0.00018126146118760137  яб -> 0.00042520211923446677  яв -> 0.0019142565527288744  яг -> 0.0004497655882739081

Entropy is: 4.161762410713714
Bigram for no cross, no spaces

```

H = 4.16176

таб. 4

```
Bigram for no cross, no spaces |
aa -> 0.000215141925909864 ab -> 0.0016787846344620096 ав -> 0.004372293785617 аг -> 0.0010909559077634048 ад ->
ба -> 0.001560202470574683 бб -> 2.1175386408451182e-05 бв -> 4.150375736056432e-05 бг -> 2.2022401864789228e-05
ва -> 0.006969243174749453 вб -> 0.0001507687512281724 вв -> 0.0003421942443605711 вг -> 0.0002473285132507098 в,
га -> 0.0014467023994253847 гб -> 5.3361973749296974e-05 гв -> 9.063065382817105e-05 гг -> 8.470154563380473e-06
да -> 0.004650961870752217 дб -> 0.00011265305569296029 дв -> 0.0012324074889718587 дг -> 6.522019013802965e-05 ,
еа -> 0.0002168359568225401 еб -> 0.0024851433488958306 ев -> 0.0034490469382085285 ег -> 0.004447678161231086 е,
жа -> 0.0016034002588479234 жб -> 1.4399262757746803e-05 жв -> 1.6093293670422896e-05 жг -> 2.795151005915556e-05
за -> 0.006252668098687465 зб -> 0.0002007426631521172 зв -> 0.0008741199509408648 зг -> 0.00043790699092677044 :
иа -> 0.0003786159089831071 иб -> 0.0011341536960366454 ив -> 0.005770716304031116 иг -> 0.0010867208304817147 и,
йа -> 0.00014399262757746803 йб -> 0.00024478746688169563 йв -> 0.0007131870142366357 йг -> 0.0002253061113859205 !
ка -> 0.009034266857301613 кб -> 0.0003498173834676135 кв -> 0.0006013809740000136 кг -> 0.00011011200932394615 .
ла -> 0.0047983425601550375 лб -> 0.00025918672963944246 лв -> 0.000663213102312691 лг -> 0.0003820039708084593 .
ма -> 0.0033287707434085257 мб -> 0.0003692987389633886 мв -> 0.0008080527453464971 мг -> 0.0002490225441633859 !
на -> 0.01120347344098335 нб -> 0.00018549638493803235 нв -> 0.00041673160451831924 нг -> 0.00020158967860845525
оа -> 0.00021260087954084986 об -> 0.005327727220366317 ов -> 0.00865395691740583 ог -> 0.005604701274588859 од
па -> 0.002096363254436667 пб -> 3.3880618253521892e-06 пв -> 6.7761236507043785e-06 пг -> 4.2350772816902365e-06
ра -> 0.008686990520203013 рб -> 0.0005344667529493079 рв -> 0.0006953996896535368 рг -> 0.0001880374313070465 р,
са -> 0.0021937700319155425 сб -> 0.0002058247558901455 св -> 0.0018134600920197592 сг -> 0.00017872026128732798
та -> 0.007836587002039614 тб -> 0.00029476137880564044 тв -> 0.002851054026033867 тг -> 0.0003167837806704297 т,
уа -> 0.00017448518400563773 уб -> 0.0008376982863183287 ув -> 0.0012832284163521415 уг -> 0.001487359141329611 !
фа -> 0.00029137331698028825 фб -> 4.2350772816902365e-06 фв -> 3.388061825352189e-05 фг -> 3.3880618253521892e-05
ха -> 0.00063272905458845213 хб -> 0.00016008592124789093 хв -> 0.000663213102312691 хг -> 9.740677747887544e-05 :
ца -> 0.0006428847313605778 цб -> 3.3880618253521892e-06 цв -> 0.00016770906035493336 цг -> 1.0164185476056568e-05
ча -> 0.0023064230876085026 чб -> 7.623139107042425e-06 чв -> 3.1339571884507746e-05 чг -> 1.1858216388732662e-05
ша -> 0.0011680343142901671 шб -> 8.470154563380473e-07 шв -> 6.860825196338183e-05 шг -> 8.470154563380473e-07 !
ща -> 0.0005488660157070546 щб -> 8.470154563380473e-07 щв -> 0.0 щг -> 0.0 шд -> 0.0 ще -> 0.00177026230374651!
ья -> 6.352615922535354e-05 ьб -> 0.0004243547436253617 ьв -> 0.0021099155017380755 ьг -> 0.0003421942443605711 !
ьа -> 0.00016008592124789093 ьб -> 0.0003210188579521199 ьв -> 0.0011138253250845321 ьг -> 0.0002219180495605684
за -> 0.0 зб -> 0.0 зв -> 1.6940309126760946e-06 зг -> 1.3552247301408757e-05 зд -> 6.7761236507043785e-06 зе -> .
юа -> 3.9809726447888224e-05 юб -> 0.0005166794283662089 юв -> 0.0002278471577549347 юг -> 0.00011773514843098857 :
яа -> 0.0001812613076563421 яб -> 0.00042520175908169973 яв -> 0.001914254931323987 яг -> 0.0004497652073155031 !
Entropy is: 4.161759496628871
```

H = 4.16175

Для біграм з урахуванням пробілів наведено фрагменти з отриманого файлу after.txt: з перетинами (таб. 6) та без (таб. 5)

таб. 5

```
Bigram cross, with spaces
aa -> 2.335504408795368e-05 ab -> 0.0010177137393477997 ав -> 0.002600194908458843 аг -> 0.0005973229457646336 ад -> 0.001!
ба -> 0.0013080051828381473 бб -> 1.3446843565791513e-05 бв -> 2.6893687131583025e-05 бг -> 1.62775800069499e-05 бд -> 1.!
ва -> 0.005771526804159463 вб -> 9.200471913436299e-06 вв -> 4.034053069737454e-05 вг -> 1.4154572174517383e-05 вд -> 0.00!
га -> 0.0012045540920514292 гб -> 7.077286087258692e-07 гв -> 1.769321521814673e-05 гг -> 1.4154572174517383e-06 гд -> 0.0!
да -> 0.0038726909469479556 дб -> 6.86496750464093e-05 дв -> 0.0009398635923879542 дг -> 1.2739114957065644e-05 дд -> 2.54!
еа -> 5.0956459828262575e-05 еб -> 0.001314959755012665 ев -> 0.001345392085187877 ег -> 0.0033666649917089594 ед -> 0.002!
жа -> 0.0013383147991006185 жб -> 7.0772860872586915e-06 жв -> 4.246371652355215e-06 жг -> 2.2647315479227813e-05 жд -> 0.!
за -> 0.005212421203266026 зб -> 0.0001372993500928186 зв -> 0.0006383712050707339 зг -> 0.0003340479033186102 зд -> 0.000!
иа -> 0.0001372993500928186 иб -> 0.000394912563669035 ив -> 0.0032024719544845575 иг -> 0.000551320586197452 ид -> 0.0016!
йа -> 7.0772860872586915e-06 йб -> 8.49274330471043e-06 йв -> 1.2031386348339775e-05 йг -> 6.369557478532822e-06 йд -> 0.0!
ка -> 0.0074927227805807765 кб -> 3.5386430436293458e-06 кв -> 0.0001330529784404634 кг -> 3.5386430436293458e-06 кд -> 1.!
ла -> 0.003927186049819848 лб -> 4.671008817590736e-05 лв -> 1.9108672435598466e-05 лг -> 0.00015499256531096534 лд -> 3.6!
ма -> 0.002686537798723399 мб -> 5.307964565444018e-05 мв -> 2.477050130540542e-05 мг -> 1.3446843565791513e-05 мд -> 1.13!
на -> 0.009334940349094214 нб -> 3.5386430436293458e-06 нв -> 2.6893687131583025e-05 нг -> 0.00010969793435250971 нд -> 0.!
оа -> 9.908200522162167e-06 об -> 0.0035690753738045577 ов -> 0.005379445154925331 ог -> 0.004336960914272126 од -> 0.0043!
па -> 0.0017495051207703485 пб -> 0.0 пв -> 0.0 пг -> 0.0 пд -> 0.0 пе -> 0.0023107339074899627 пж -> 0.0 пз -> 7.07728!
ра -> 0.0072478486819616255 рб -> 0.00042746807967042496 рв -> 0.0005201805274135138 рг -> 0.0001422534503538997 рд -> 0.0!
са -> 0.0018032924950335146 сб -> 7.78501469598456e-05 св -> 0.001311421119690355 сг -> 4.317144513227801e-05 сд -> 0.000!
та -> 0.006443161253840313 тб -> 4.034053069737454e-05 тв -> 0.001655377215809808 тг -> 2.0524129653050203e-05 тд -> 0.000!
уа -> 6.298784617660235e-05 уб -> 0.0005711369872417764 ув -> 0.0005923688455035525 уг -> 0.001140858517266101 уд -> 0.001!
фа -> 0.0002413354557552136 фб -> 0.0 фв -> 1.4154572174517383e-06 фг -> 0.0 фд -> 7.077286087258692e-07 фе -> 0.0002094!
ха -> 0.0004763013536725099 хб -> 0.0 хв -> 0.00023425816948826266 хг -> 4.954100261081084e-06 хд -> 5.661828869806953e-06
ца -> 0.0005279655421094984 цб -> 0.0 цв -> 0.00011394430600486493 цг -> 0.0 цд -> 0.0 це -> 0.0005725524444592281 цж -> 0.0!
ча -> 0.001925021815734364 чб -> 0.0 чв -> 2.1231858261776075e-06 чг -> 2.1231858261776075e-06 чд -> 0.0 че -> 0.00384367!
ша -> 0.0009745422942155217 шб -> 0.0 шв -> 4.671008817590736e-05 шг -> 0.0 шд -> 0.0 ше -> 0.0019391763879088814 шж -> !
ща -> 0.0004586081384543632 щб -> 0.0 щв -> 0.0 щг -> 0.0 шд -> 0.0 ще -> 0.0014784450636283406 щж -> 0.0 щз -> 0.0 щи
ья -> 7.077286087258692e-07 ьб -> 0.00022859634061845571 ьв -> 0.0013708703151020084 ьг -> 0.0002130263112264866 ьд -> 0.0!
ьа -> 5.661828869806953e-06 ьб -> 3.6094159045019325e-05 ьв -> 4.1048259306100406e-05 ьг -> 3.467870182756759e-05 ьд -> 2.!
за -> 0.0 зб -> 0.0 зв -> 1.4154572174517383e-06 зг -> 1.1323657739613906e-05 зд -> 5.661828869806953e-06 зе -> 0.0 эж -> !
юа -> 2.1231858261776075e-06 юб -> 0.0003666034193200002 юв -> 9.200471913436299e-06 юг -> 5.661828869806953e-06 юд -> 0.0!
яа -> 6.369557478532822e-06 яб -> 1.769321521814673e-05 яв -> 0.0003121083164481083 яг -> 0.00013022206400555993 яд -> 0.0!
а -> 0.0017452587491179933 б -> 0.005228698961266721 в -> 0.013657746691191823 г -> 0.0034579619822345965 д -> 0.0066!
Entropy is: 3.9982473110824768
```

$$H = 3.99824$$

таб. 6

Bigram for no cross, with spaces  
аа -> 2.3355027558932518e-05    аб -> 0.0010177130190831807    ав -> 0.0026001930682278203    аг -> 0.00059732252302239!  
ба -> 0.0013008042622217567    бб -> 1.344683404908236e-05    бв -> 2.689366809816472e-05    бг -> 1.627774648046812e-0!  
ва -> 0.005771522719487718    вб -> 9.20046540200372e-06    вв -> 4.034050214724708e-05    вг -> 1.4154562156928799e-05  
га -> 0.0012045532395546408    гб -> 7.0772810784644e-07    гв -> 1.7693202696160998e-05    гг -> 1.41545621569288e-06  
да -> 0.0038726882061357196    дб -> 6.864962646110468e-05    дв -> 0.0009398629272200723    дг -> 1.273910594123592e-0!  
еа -> 5.095642376494368e-05    еб -> 0.0013149588243786855    ев -> 0.0013453911330160823    ег -> 0.003366662609025515  
жа -> 0.001338313851937618    жб -> 7.0772810784643995e-06    жв -> 4.24636864707864e-06    жг -> 2.264729945108608e-05  
за -> 0.00521241751428903    зб -> 0.00013729925292220937    зв -> 0.0006383707532774889    зг -> 0.0003340476669035197  
иа -> 0.00013729925292220937    иб -> 0.0003949122841783135    ив -> 0.003202469688005141    иг -> 0.000551320196012376!  
йа -> 7.0772810784643995e-06    йб -> 8.49273729415728e-06    йв -> 1.203137783338948e-05    йг -> 6.36955297061796e-06  
ка -> 0.00749271747777026    кб -> 3.5386405392321998e-06    кв -> 0.00013305288427513073    кг -> 3.5386405392321998e-(!  
ла -> 0.003927183270439895    лб -> 4.6710055117865037e-05    лв -> 1.910865891185388e-05    лг -> 0.000154992455618370!  
ма -> 0.0026865358973850863    мб -> 5.3079608088483e-05    мв -> 2.47704837746254e-05    мг -> 1.344683404908236e-05    !  
на -> 0.009334933742494543    нб -> 3.5386405392321998e-06    нв -> 2.689366809816472e-05    нг -> 0.000109697856716198!  
оа -> 9.90819350985016e-06    об -> 0.003569072847869597    ов -> 0.005379441347740791    ог -> 0.0043369578448829845    (!  
па -> 0.0017495038825963997    пб -> 0.0    пв -> 0.0    пг -> 0.0    пд -> 0.0    пе -> 0.0023107322721186265    пж -> 0.0    !  
ра -> 0.007247843552455392    рб -> 0.0004274677713924977    рв -> 0.0005201801592671334    рг -> 0.000142253349677134!  
са -> 0.001803291218792729    сб -> 7.78500918631084e-05    св -> 0.0013114201838394532    сг -> 4.317141457863284e-05  
та -> 0.00644315669383399    тб -> 4.034050214724708e-05    тв -> 0.001655376044252823    тг -> 2.052411512754676e-05    -  
уа -> 6.298780159833316e-05    уб -> 0.000571136583032077    ув -> 0.0005923684262674702    уг -> 0.0011408577098484612  
фа -> 0.00024133528477563603    фб -> 0.0    фв -> 1.41545621569288e-06    фг -> 0.0    фд -> 7.0772810784644e-07    фе -> (!  
ха -> 0.0004763010165806541    хб -> 0.0    хв -> 0.00023425800369717163    хг -> 4.95409675492508e-06    хд -> 5.6618248!  
ца -> 0.0005279651684534443    цб -> 0.0    цв -> 0.00011394422536327683    цг -> 0.0    цд -> 0.0    це -> 0.0005725520392!  
ча -> 0.0019250204533423168    чб -> 0.0    чв -> 2.12318432353932e-06    чг -> 2.12318432353932e-06    чд -> 0.0    че -> (!  
ша -> 0.0009745416045045479    шб -> 0.0    шв -> 4.6710055117865037e-05    шг -> 0.0    шд -> 0.0    ше -> 0.0019391750154!  
ща -> 0.0004586078138844931    щб -> 0.0    щв -> 0.0    щг -> 0.0    щд -> 0.0    ще -> 0.0014784440172912131    щж -> 0.0    !  
ья -> 7.0772810784644e-07    ыб -> 0.0002285961788344001    ыв -> 0.0013708693448985542    ыг -> 0.00021302616046177844  
ья -> 5.66182486277152e-06    ьб -> 3.6094133500168436e-05    ьв -> 4.104823025509352e-05    ьг -> 3.467867728447556e-0!  
эа -> 0.0    эб -> 0.0    эв -> 1.41545621569288e-06    эг -> 1.132364972554304e-05    эд -> 5.66182486277152e-06    эе -> (!  
юа -> 2.12318432353932e-06    юб -> 0.0003666031598644559    юв -> 9.20046540200372e-06    юг -> 5.66182486277152e-06    !  
яа -> 6.36955297061796e-06    яб -> 1.7693202696160998e-05    яв -> 0.00031210809556028    яг -> 0.00013022197184374496  
а -> 0.001745257513949321    б -> 0.005228695260769498    в -> 0.0136577370252206    г -> 0.0034579595349377057    д

Entropy is: 3.998244991928201

$$H = 3.9982$$

За допомогою CoolPinkProgram було оцінено значення  $H^{10}$ ,  $H^{20}$ ,  $H^{30}$ .



Произвольная часть текста:  
ставить\_на\_первое\_место\_самого\_себя\_эгоизм\_никогда\_и\_нигде\_не\_считался\_похв

Использованные буквы:

Порядок n-граммы:  
5 символов  
10 символов  
15 символов  
20 символов  
25 символов  
30 символов  
35 символов  
40 символов  
45 символов  
50 символов

Введенный символ: а

Символ по счету: 1

Номер эксперимента: 50

Неравенство для энтропии:  
 $1,94072451102478 < H < 2,74791993994523$

Двоичная таблица угаданных символов:

01000000000000000000000000000000
00000000000000000000000000000000
00000100000000000000000000000000
10000000000000000000000000000000
10000000000000000000000000000000

Поле ввода символов:  
а

Продолжить Другой

Вероятности:

q[1] = 0,46
q[2] = 0,16
q[3] = 0,04
q[4] = 0,08
q[5] = 0,06
q[6] = 0,02
q[7] = 0
q[8] = 0,04
q[9] = 0
q[10] = 0,02
q[11] = 0
q[12] = 0,02
q[13] = 0
q[14] = 0
q[15] = 0
q[16] = 0
q[17] = 0
q[18] = 0,02
q[19] = 0,02
q[20] = 0
q[21] = 0,02
q[22] = 0
q[23] = 0
q[24] = 0
q[25] = 0
q[26] = 0,02
q[27] = 0
q[28] = 0
q[29] = 0
q[30] = 0,02
q[31] = 0
q[32] = 0

Строка состояния:  
Вы угадали. Для продолжения опыта нажмите "Продолжить", или "Другой" для выбора другого порядка

$$H^{10} = \frac{1,94 + 2,75}{2} = 2,345$$

Произвольная часть текста:  
ему\_закону\_человеческой\_природы\_или\_нарушать\_его\_ту\_же\_идею\_можно\_выразить\_

Использованные буквы:

Порядок n-граммы:  
5 символов  
10 символов  
15 символов  
20 символов  
25 символов  
30 символов  
35 символов  
40 символов  
45 символов  
50 символов

Введенный символ: с

Символ по счету: 1

Номер эксперимента: 51

Неравенство для энтропии:  
 $0,708623630463532 < H < 1,1130408664833$

Двоичная таблица угаданных символов:

10000000000000000000000000000000
10000000000000000000000000000000
10000000000000000000000000000000
10000000000000000000000000000000
10000000000000000000000000000000

Поле ввода символов:  
с

Продолжить Другой

Вероятности:

q[1] = 0,7843137
q[2] = 0,1372549
q[3] = 0
q[4] = 0
q[5] = 0
q[6] = 0
q[7] = 0
q[8] = 0
q[9] = 0
q[10] = 0,019607
q[11] = 0
q[12] = 0
q[13] = 0
q[14] = 0
q[15] = 0
q[16] = 0,019607
q[17] = 0
q[18] = 0
q[19] = 0
q[20] = 0,019607
q[21] = 0
q[22] = 0
q[23] = 0
q[24] = 0
q[25] = 0
q[26] = 0
q[27] = 0
q[28] = 0
q[29] = 0
q[30] = 0,019607
q[31] = 0
q[32] = 0

Строка состояния:  
Вы угадали. Для продолжения опыта нажмите "Продолжить", или "Другой" для выбора другого порядка

$$H^{20} = \frac{0,71 + 1,11}{2} = 0,91$$

[illegible]

$$H^{30} = \frac{1,36+2,07}{2} = 1,715$$

$N_0$  для заданого алфавіту –  $\log_2 32=5$ , тоді:

$$\begin{aligned} 1) \quad R &= 1 - \frac{2,345}{5} = 0,531 \\ 2) \quad R &= 1 - \frac{0,91}{5} = 0,818 \\ 3) \quad R &= 1 - \frac{1,715}{5} = 0,65 \end{aligned}$$

Середнє значення  $R$  надлишковості російської мови – 0,666, але кількість проведених експериментів ніяк не прямує до нескінченності, тому можуть виникати значні похибки.

**Висновки:** після виконання цієї лабораторної роботи в нас з'явилося більш чітке уявлення про поняття ентропії не лише у вигляді формули, а на конкретному прикладі мови. Підтвердили, що не має істотної різниці у розподілі частот при підрахунку їх для біграм з перетином і без, що свідчить про коректність програми.

І з першої, і з другої частини роботи побачили, як часто зустрічається пробіл. Друга частина показала, що надлишковість російської мови – приблизно 66%, що свідчить про повторюваність значної кількості інформації, але є цілком природнім для усіх мов.

Також було отримано навички роботи з мовою Python та системою Git.